



1645

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/841,963A

DATE: 02/13/2002
TIME: 15:16:37

Input Set : A:\10545-015.txt
Output Set: N:\CRF3\02132002\I841963A.raw

Does Not Comply
Corrected Diskette Needed

3 <110> APPLICANT: Watson, Dennis K.
4 Papas, Takis S. (Deceased)
5 Papas, Tula C. (Legal Representative)
7 <120> TITLE OF INVENTION: Methods and compositions for the diagnosis and treatment of cancer
8 based on transcription factor ETS2
10 <130> FILE REFERENCE: 10545-015-999
12 <140> CURRENT APPLICATION NUMBER: 09/841,963A
13 <141> CURRENT FILING DATE: 2001-04-25
15 <150> PRIOR APPLICATION NUMBER: PCT/US99/27805
16 <151> PRIOR FILING DATE: 1999-11-23
18 <150> PRIOR APPLICATION NUMBER: 06/109,850 *← please edit*
19 <151> PRIOR FILING DATE: 1998-11-25
21 <160> NUMBER OF SEQ ID NOS: 6
23 <170> SOFTWARE: PatentIn version 3.0
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 1894
27 <212> TYPE: DNA
28 <213> ORGANISM: Homo sapiens
30 <220> FEATURE:
31 <221> NAME/KEY: CDS
32 <222> LOCATION: (416)..(1423)
34 <400> SEQUENCE: 1
35 gtctgacttc ctcccagcac attcctgcac tctgccgtgt ccacactgcc ccacagaccc 60
36 agtcctccaa gcctgctgcc agtcctcgc aagccctca ggttggccct tgccacgggt 120
37 ccagcaggca gcccgggt gggggtaggg gactccctac aggacacgcg ccctgagacc 180
38 gccaccagcc acccctttag ggtggccagg ccccccagtgg ccaaccttag tgctgcctct 240
39 gccaccagcc ctgctggccc ctggttccgc tggcccccga gatgcctggc tgagacacgc 300
40 cagtggcctc agctgcccac acctttccc ggcccttggg gttggactg cagcagacag 360
41 ctccctgggc accaggcagc taacagacac agccggcagc ccaaacagca gcggc atg 418
42 Met
43 1
44 ggc agc gcc agc ccg ggt ctg agc agc gta tcc ccc agc cac ctc ctg 466
45 Gly Ser Ala Ser Pro Gly Leu Ser Ser Val Ser Pro Ser His Leu Leu
46 5 10 15
47 ctg ccc ccc gac acg gtg tcg cgg aca ggc ttg gag aag gcg gca gcg 514
48 Leu Pro Pro Asp Thr Val Ser Arg Thr Gly Leu Glu Lys Ala Ala Ala
49 20 25 30
50 ggg gca gtg ggt ctc gag aga cgg gac tgg agt ccc agt cca ccc gcc 562
51 Gly Ala Val Gly Leu Glu Arg Arg Asp Trp Ser Pro Ser Pro Pro Ala
52 35 40 45
53 acg ccc gag cag ggc ctg tcc gcc ttc gac ctc tcc tac ttt gac atg 610
54 Thr Pro Glu Gln Gly Leu Ser Ala Phe Asp Leu Ser Tyr Phe Asp Met
55 50 55 60 65

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/841,963A

DATE: 02/13/2002

TIME: 15:16:37

Input Set: A:\10545-015.txt

Output Set: N:\CRF3\02132002\I841963A.raw

56	ctg tac cct gag gac agc agc tgg gca aag gcc cct ggg gcc agc	658
57	Leu Tyr Pro Glu Asp Ser Ser Trp Ala Ala Lys Ala Pro Gly Ala Ser	
58	70 75 80	
59	agt cgg gag gag cca cct gac cag cct gag cag tgc ccg gtc att gac	706
60	Ser Arg Glu Glu Pro Pro Asp Gln Pro Glu Gln Cys Pro Val Ile Asp	
61	85 90 95	
62	agc caa gcc cca gcg ggc acc ctg gac ttg gtg ccc ggc ggg ctg acc	754
63	Ser Gln Ala Pro Ala Gly Thr Leu Asp Leu Val Pro Gly Gly Leu Thr	
64	100 105 110	
65	ttg gag gag cac tcg ctg qag cag gtg cag tcc atq gtg gtg ggc gaa	802
66	Leu Glu Glu His Ser Leu Glu Gln Val Gln Ser Met Val Val Gly Glu	
67	115 120 125	
68	gtg ctc aag gac atc gag acg gcc tgc aag ctg ctc aac atc acc gca	850
69	Val Leu Lys Asp Ile Glu Thr Ala Cys Lys Leu Leu Asn Ile Thr Ala	
70	130 135 140 145	
71	gat ccc atg gac tgg agc ccc agc aat gtg cag aag tgg ctc ctg tgg	898
72	Asp Pro Met Asp Trp Ser Pro Ser Asn Val Gln Lys Trp Leu Leu Trp	
73	150 155 160	
74	aca gag cac caa tac cgg ctg ccc atg ggc aag gcc ttc cag gag	946
75	Thr Glu His Gln Tyr Arg Leu Pro Pro Met Gly Lys Ala Phe Gln Glu	
76	165 170 175	
77	ctg gcg ggc aag gag ctg tgc gcc atg tcg gag gag cag ttc cgc cag	994
78	Leu Ala Gly Lys Glu Leu Cys Ala Met Ser Glu Glu Gln Phe Arg Gln	
79	180 185 190	
80	cgc tcg ccc ctg ggt ggt gat gtg ctg cac gcc cac ctg gac atc tgg	1042
81	Arg Ser Pro Leu Gly Gly Asp Val Leu His Ala His Leu Asp Ile Trp	
82	195 200 205	
83	aag tca gcg gcc tgg atg aaa gag cgg act tca cct ggg gcg att cac	1090
84	Lys Ser Ala Ala Trp Met Lys Glu Arg Thr Ser Pro Gly Ala Ile His	
85	210 215 220 225	
86	tac tgt gcc tcg acc agt gag gag agc tgg acc gac agc gag gtg gac	1138
87	Tyr Cys Ala Ser Thr Ser Glu Glu Ser Trp Thr Asp Ser Glu Val Asp	
88	230 235 240	
89	tca tca tgc tcc ggg cag ccc atc cac ctg tgg cag ttc ctc aag gag	1186
90	Ser Ser Cys Ser Gly Gln Pro Ile His Leu Trp Gln Phe Leu Lys Glu	
91	245 250 255	
92	ttg cta ctc aag ccc cac agc tat ggc cgc ttc att agg tgg ctc aac	1234
93	Leu Leu Leu Lys Pro His Ser Tyr Gly Arg Phe Ile Arg Trp Leu Asn	
94	260 265 270	
95	aag gag aag ggc atc ttc aaa att gag gac tca gcc cag gtg gcc cgg	1282
96	Lys Glu Lys Gly Ile Phe Lys Ile Glu Asp Ser Ala Gln Val Ala Arg	
97	275 280 285	
98	ctg tgg ggc atc cgc aag aac cgt ccc gcc atg aac tgc gac aag ctg	1330
99	Leu Trp Gly Ile Arg Lys Asn Arg Pro Ala Met Asn Cys Asp Lys Leu	
100	290 295 300 305	
101	agc cgc tcc atc cgc cag tct tac aag aag ggc atc atc cgg aag cca	1378
102	Ser Arg Ser Ile Arg Gln Ser Tyr Lys Lys Gly Ile Ile Arg Lys Pro	
103	310 315 320	
104	gac atc tcc cag cgc ctc gtc tac cag ttc gtg cac ccc atc tga	1423

RAW SEQUENCE LISTING DATE: 02/13/2002
 PATENT APPLICATION: US/09/841,963A TIME: 15:16:37

Input Set : A:\10545-015.txt
 Output Set: N:\CRF3\02132002\I841963A.raw

105 Asp Ile Ser Gln Arg Leu Val Tyr Gln Phe Val His Pro Ile
 106 325 330 335
 107 gtgcctggcc cagggcctga aaccgcctc caggggcctc tctcctgcct gccctgcctc 1483
 108 agccaggccc tgagatgggg gaaaacgggc agtgtgctct gctgctctga cttccagag 1543
 109 cccaagggtca gggagggggca accaactgcc ccagggggat atgggtcctc tggggccttc 1603
 110 gggaccatgg ggcagggggtg ctccctcctc aggcccagct gctcccctgg aggacagagg 1663
 111 gagacagggc tgctcccaa cacctgcctc tgaccccagc atttccagag cagagcctac 1723
 112 agaagggcag tgactcgaca aaggccacag gcagttcagg cctctcttg ctccatcccc 1783
 113 ctgcctccca ttctgcacca cacctggcat ggtcaggga gacatctgca cccctcagtt 1843
 114 gggcagccag gagtgcccccc gggatggat aataaagata ctagagaact g 1894
 116 <210> SEQ ID NO: 2
 117 <211> LENGTH: 335
 118 <212> TYPE: PRT
 119 <213> ORGANISM: Homo sapiens
 121 <400> SEQUENCE: 2
 122 Met Gly Ser Ala Ser Pro Gly Leu Ser Ser Val Ser Pro Ser His Leu
 123 1 5 10 15
 125 Leu Leu Pro Pro Asp Thr Val Ser Arg Thr Gly Leu Glu Lys Ala Ala
 126 20 25 30
 128 Ala Gly Ala Val Gly Leu Glu Arg Arg Asp Trp Ser Pro Ser Pro Pro
 129 35 40 45
 131 Ala Thr Pro Glu Gln Gly Leu Ser Ala Phe Asp Leu Ser Tyr Phe Asp
 132 50 55 60
 134 Met Leu Tyr Pro Glu Asp Ser Ser Trp Ala Ala Lys Ala Pro Gly Ala
 135 65 70 75 80
 137 Ser Ser Arg Glu Glu Pro Pro Asp Gln Pro Glu Gln Cys Pro Val Ile
 138 85 90 95
 140 Asp Ser Gln Ala Pro Ala Gly Thr Leu Asp Leu Val Pro Gly Gly Leu
 141 100 105 110
 143 Thr Leu Glu Glu His Ser Leu Glu Gln Val Gln Ser Met Val Val Gly
 144 115 120 125
 146 Glu Val Leu Lys Asp Ile Glu Thr Ala Cys Lys Leu Leu Asn Ile Thr
 147 130 135 140
 149 Ala Asp Pro Met Asp Trp Ser Pro Ser Asn Val Gln Lys Trp Leu Leu
 150 145 150 155 160
 152 Trp Thr Glu His Gln Tyr Arg Leu Pro Pro Met Gly Lys Ala Phe Gln
 153 165 170 175
 155 Glu Leu Ala Gly Lys Glu Leu Cys Ala Met Ser Glu Glu Gln Phe Arg
 156 180 185 190
 158 Gln Arg Ser Pro Leu Gly Gly Asp Val Leu His Ala His Leu Asp Ile
 159 195 200 205
 161 Trp Lys Ser Ala Ala Trp Met Lys Glu Arg Thr Ser Pro Gly Ala Ile
 162 210 215 220
 164 His Tyr Cys Ala Ser Thr Ser Glu Glu Ser Trp Thr Asp Ser Glu Val
 165 225 230 235 240
 167 Asp Ser Ser Cys Ser Gly Gln Pro Ile His Leu Trp Gln Phe Leu Lys
 168 245 250 255
 170 Glu Leu Leu Leu Lys Pro His Ser Tyr Gly Arg Phe Ile Arg Trp Leu
 171 260 265 270

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/841,963A

DATE: 02/13/2002
TIME: 15:16:37

Input Set : A:\10545-015.txt
Output Set: N:\CRF3\02132002\I841963A.raw

173 Asn Lys Glu Lys Gly Ile Phe Lys Ile Glu Asp Ser Ala Gln Val Ala
174 275 280 285
176 Arg Leu Trp Gly Ile Arg Lys Asn Arg Pro Ala Met Asn Cys Asp Lys
177 290 295 300
179 Leu Ser Arg Ser Ile Arg Gln Ser Tyr Lys Lys Gly Ile Ile Arg Lys
180 305 310 315 320
182 Pro Asp Ile Ser Gln Arg Leu Val Tyr Gln Phe Val His Pro Ile
183 325 330 335
185 <210> SEQ ID NO: 3
186 <211> LENGTH: 1704
187 <212> TYPE: DNA
188 <213> ORGANISM: Mus sp.
190 <220> FEATURE:
191 <221> NAME/KEY: CDS
192 <222> LOCATION: (362)..(1336)
194 <400> SEQUENCE: 3
195 tgtccgctct gcctccaca cctagcaccc cagcccgctg ctgccccgg gagaacccccc 60
196 agcttggggc ttgtcatggt gccagcagg ggcctgagc ttctgacagg ggcctgccta 120
197 tagacctgca ggcctgagc ctcagactca cactcaagg gcaagaggcc ctggggccc 180
198 acctaagagc cacctctgtc cccagccctg ctgccccact gatgtctgac tgagacccag 240
199 cagtgaccct gagctgcctg cccactgctt cctctggtc cctgagggtt gctctgcccga 300
200 ggacggacga ctcttctgaa gcaggcggtt aacgaaagca gccccaaagcc tccaccgcag 360
201 c atg ggc agt gcc agc cca ggc ctg agc aac gtg tcc ccc ggt tgc ctg 409
202 Met Gly Ser Ala Ser Pro Gly Leu Ser Asn Val Ser Pro Gly Cys Leu
203 1 5 10 15
204 cta ctg ttc cca gat gtg gca cca cga aca ggg acg gag aag gca gca 457
205 Leu Leu Phe Pro Asp Val Ala Pro Arg Thr Gly Thr Glu Lys Ala Ala
206 20 25 30
207 tca gga gca atg ggc cct gag aag cag gaa tgg agt cct agt cca ccc 505
208 Ser Gly Ala Met Gly Pro Glu Lys Gln Glu Trp Ser Pro Ser Pro Pro
209 35 40 45
210 gcc acc cct gag cag ggc ctg tct gct ttc tac ctc tct tac ttt aac 553
211 Ala Thr Pro Glu Gln Gly Leu Ser Ala Phe Tyr Leu Ser Tyr Phe Asn
212 50 55 60
213 atg tat ccc gac gat agc agc tgg gtc gcc aaa gtc ccc gag gcc cgt 601
214 Met Tyr Pro Asp Asp Ser Ser Trp Val Ala Lys Val Pro Glu Ala Arg
215 65 70 75 80
216 gcc ggg gag gac cac ccg gag gag ccc gag cag tgt ccc gtc att gac 649
217 Ala Gly Glu Asp His Pro Glu Glu Pro Glu Gln Cys Pro Val Ile Asp
218 85 90 95
219 agc cag gcc tct ggg agc acg ttg gat gag cac tcg cta gag cag gtg 697
220 Ser Gln Ala Ser Gly Ser Thr Leu Asp Glu His Ser Leu Glu Gln Val
221 100 105 110
222 caa tcg atg gtt gtg ggc gag gtc ctg aaa gat att gag acg gcc tgc 745
223 Gln Ser Met Val Val Gly Glu Val Leu Lys Asp Ile Glu Thr Ala Cys
224 115 120 125
225 aag ctt gtg aac atc aca gca gac cct ggg gac tgg agc cct ggt aac 793
226 Lys Leu Val Asn Ile Thr Ala Asp Pro Gly Asp Trp Ser Pro Gly Asn
227 130 135 140

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/841,963A

DATE: 02/13/2002
TIME: 15:16:37

Input Set : A:\10545-015.txt
Output Set: N:\CRF3\02132002\I841963A.raw

228	gtg	cag	aag	tgg	ctt	tta	tgg	aca	gaa	cac	cag	tac	cg	ctg	cct	cca	841
229	Val	Gln	Lys	Trp	Leu	Leu	Trp	Thr	Glu	His	Gln	Tyr	Arg	Leu	Pro	Pro	
230	145				150				155					160			
231	gca	ggc	aag	gcc	ttc	cag	gag	ctg	ggc	ggt	aag	gag	ctg	tgc	gcc	atg	889
232	Ala	Gly	Lys	Ala	Phe	Gln	Glu	Leu	Gly	Gly	Lys	Glu	Leu	Cys	Ala	Met	
233					165				170					175			
234	tcc	gag	gaa	cag	ttc	cgt	cag	cgt	gca	ccc	ttg	ggt	ggg	gat	gta	ctg	937
235	Ser	Glu	Glu	Gln	Phe	Arg	Gln	Arg	Ala	Pro	Leu	Gly	Gly	Asp	Val	Leu	
236					180				185					190			
237	cat	gcc	cac	ctg	gac	atc	tgg	aag	tca	gcg	gcc	tgg	atg	aag	gag	agg	985
238	His	Ala	His	Leu	Asp	Ile	Trp	Lys	Ser	Ala	Ala	Trp	Met	Lys	Glu	Arg	
239					195				200					205			
240	acc	tcg	cct	ggg	acc	ctt	cac	tac	tgc	gcc	tcc	acc	agc	gag	gac	ggc	1033
241	Thr	Ser	Pro	Gly	Thr	Leu	His	Tyr	Cys	Ala	Ser	Thr	Ser	Glu	Asp	Gly	
242					210				215					220			
243	tgg	acg	gat	ggt	gag	gtg	gac	tcg	tcg	tcg	tcg	ggg	cag	ccc	att	cac	1081
244	Trp	Thr	Asp	Gly	Glu	Val	Asp	Ser	Ser	Cys	Ser	Gly	Gln	Pro	Ile	His	
245	225				230				235					240			
246	ctg	tgg	cag	ttc	ctg	aaa	gaa	ctg	ctg	ctc	aag	ccc	cac	agc	tat	ggc	1129
247	Leu	Trp	Gln	Phe	Leu	Lys	Glu	Leu	Leu	Leu	Lys	Pro	His	Ser	Tyr	Gly	
248					245				250					255			
249	cgc	ttc	atc	cg	tgg	gg	gtg	cg	aag	aac	cg	cc	ga	cc	at	tt	1177
250	Arg	Phe	Ile	Arg	Trp	Gly	Val	Arg	Lys	Asn	Arg	Pro	Ala	Met	Asn	Tyr	
251					260				265					270			
252	gat	aaa	cta	aga	agc	tcc	atc	cg	tgg	ctc	aac	aag	gag	aaa	ggc	atc	1225
253	Asp	Lys	Leu	Arg	Ser	Ser	Ile	Arg	Trp	Leu	Asn	Lys	Glu	Lys	Gly	Ile	
254					275				280					285			
255	ttc	aaa	att	gag	gac	tca	gca	cag	gtg	gcc	cg	ctc	cag	tat	ta	aag	1273
256	Phe	Lys	Ile	Glu	Asp	Ser	Ala	Gln	Val	Ala	Arg	Leu	Gln	Tyr	Tyr	Lys	
257					290				295					300			
258	aag	ggc	atg	att	cgt	aaa	ccc	gcc	atc	tct	cag	cg	ctt	gtc	ta	caa	1321
259	Lys	Gly	Met	Ile	Arg	Lys	Pro	Ala	Ile	Ser	Gln	Arg	Leu	Val	Tyr	Gln	
260					305				310					315		320	
261	ttt	gtg	cat	cca	gtc	tgagagccac	agagaccaga	ggcctacaac	ctgccccagg								1376
262	Phe	Val	His	Pro	Val												
263					325												
264	cagccact	ct	tggttggcc	tggtcctctc	tgctcact	ct	gaattcaggg	gctgctggta									1436
265	tcccagaacc	caaggtccca	gatagacagc	cactgatcta	cccatacaca	ta	tgagctctct										1496
266	gggtcataca	caggccccag	gaagatcgag	ggagctagtt	cagcacacag	gg	gactggacc										1556
267	aagtca	tc	accggacagt	gatgtcact	gtctctgtc	ct	ggccacaat	cctgtaccat									1616
268	atctggcat	tg	gtgctaagag	atgtctgtac	cctgcgttgg	ga	agccaggg	gtgcctgg									1676
269	gatggataat	aa	agacgtaa	gataact													1704
271	<210>	SEQ ID NO:	4														
272	<211>	LENGTH:	325														
273	<212>	TYPE:	PRT														
274	<213>	ORGANISM:	Mus sp.														
275	<400>	SEQUENCE:	4														
276	Met	Gly	Ser	Ala	Ser	Pro	Gly	Leu	Ser	Asn	Val	Ser	Pro	Gly	Cys	Leu	
277																	
278	1				5				10					15			

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/841,963A DATE: 02/13/2002
TIME: 15:16:38

Input Set : A:\10545-015.txt
Output Set: N:\CRF3\02132002\I841963A.raw